University of Maryland College Park

Institute for Advanced Computer Studies TR–95–38
Department of Computer Science TR–3441

On the Perturbation of Schur Complements in Positive Semidefinite Matrices*

G. W. Stewart†

March, 1995

ABSTRACT

This note gives perturbation bounds for the Schur complement of a positive definite matrix in a positive semidefinite matrix.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports`.

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742.

On the Perturbation of Schur Complements in Positive Semidefinite Matrices

G. W. Stewart

ABSTRACT

This note gives perturbation bounds for the Schur complement of a positive definite matrix in a positive semidefinite matrix.

Let A be a nonsingular matrix, and suppose A is embedded in a larger matrix

$$P = \begin{pmatrix} A & B' \\ B & C \end{pmatrix}$$

(here the prime does *not* denote a transpose). The Schur complement of A in P is the matrix

$$S = C - B'A^{-1}B.$$

Now suppose we replace P by

$$\tilde{P} = \begin{pmatrix} A+E & B'+F' \\ B+F & C+G \end{pmatrix}$$

where E, F, F', and G are presumed small. Then up to first order terms, the Schur complement of $A+E$ in $\tilde{P}$ is

$$\tilde{S} \cong S + G - F'A^{-1}B - B'A^{-1}F + B'A^{-1}EA^{-1}B.$$

It follows that in any consistent matrix norm

$$\|\tilde{S} - S\| \lesssim \|G\| + \|A^{-1}B\|\|F'\| + \|B'A^{-1}\|\|F\| + \|A^{-1}B\|\|B'A^{-1}\|\|E\|. \quad (1)$$

This bound pretty much sums up the perturbation theory of Schur complements for general matrices. The error introduced into the Schur complement by the perturbations is controled by the size of the quantities $\|A^{-1}B\|$ and $\|B'A^{-1}\|$. We can put the bounds in different forms, say by assuming that the errors E, F, F', and G are small compared to $\|A\|$, $\|B\|$, $\|B'\|$, and $\|C\|$, but the change is merely cosmetic.

It is far different when P is positive semidefinite. (We write $P \rhd 0$ to mean P is positive definite and $P \unrhd 0$ to mean that P is positive semidefinite). There are then special relations among the the submatrices. The matrix B' is B^{T}. The matrices A, C, and S are positive semidefinite. As we shall see, the spectral norm of B is bounded by the geometric mean of the norms of A and B. The purpose of this note is to show how these facts can be used to simplify the general bound (1) in such a way that it depends only on the condition of A. The chief technical difficulty is that perturbations of P can make it indefinite and destroy the nice relations between the submatrices. But as we will see, a small correction factor is sufficient to handle this problem.[1]

Specifically we are going to prove the following theorem, in which $\|\cdot\|$ denotes the spectral norm (see [2] for a definition and properties).

Theorem. Let

$$A \rhd 0$$

and

$$P \equiv \begin{pmatrix} A & B^{\mathrm{T}} \\ B & C \end{pmatrix} \unrhd 0.$$

Let

$$\tilde{P} = \begin{pmatrix} A+E & B^{\mathrm{T}}+F \\ B+F & C+G \end{pmatrix} \equiv \begin{pmatrix} \tilde{A} & \tilde{B}^{\mathrm{T}} \\ \tilde{B} & \tilde{C} \end{pmatrix},$$

where E is symmetric,

$$\|E\| \le \epsilon\|A\|, \quad \|F\| \le \epsilon\|B\|, \quad \text{and} \quad \|G\| \le \epsilon\|C\|. \tag{2}$$

Let $\kappa(A) = \|A\|\|A^{-1}\|$, and assume

$$\epsilon\kappa(A) < 1, \tag{3}$$

and

$$\epsilon\beta < 1,$$

[1] Higham [1] gives first order bounds, somewhat less refined than those here.

where

$$\beta = \frac{2\kappa^{\frac{1}{2}}(A) + \kappa(A)}{1 - \epsilon\kappa(A)}.$$

Let S be the Schur complement of A in P and $\tilde{S}$ be the Schur complement of $\tilde{A}$ in $\tilde{P}$. Then

$$\frac{\|\tilde{S} - S\|}{\|C\|} \leq \epsilon + \frac{\epsilon\beta}{1 - \epsilon\beta}. \tag{4}$$

Proof. The perturbation G of C accounts for the additive term ϵ in (4). To derive the other term, it will be convenient to assume $G = 0$. For brevity we will write κ for $\kappa(A)$.

The first step is to obtain a bound on the inverse of the positive definite square root of $\tilde{A}$. Let λ be the smallest eigenvalue of A. From standard perturbation theory we know that the smallest eigenvalue $\tilde{\lambda}$ of $\tilde{A}$ satisfies

$$\tilde{\lambda} \geq \lambda - \|E\| = \lambda(1 - \epsilon\lambda^{-1}\|A\|) = \lambda(1 - \epsilon\kappa) > 0.$$

Thus $\tilde{A}$ is positive definite and has a positive definite square root, the norm of whose inverse is

$$\|\tilde{A}^{-\frac{1}{2}}\| = \sqrt{\tilde{\lambda}^{-1}} = \frac{\lambda^{-\frac{1}{2}}}{\sqrt{1 - \epsilon\kappa}} \leq \frac{\|A^{-\frac{1}{2}}\|}{1 - \epsilon\kappa}. \tag{5}$$

Since $\tilde{A} = A + E \rhd 0$, it is nonsingular, and $\tilde{S}$ is well defined. Let δ be the least upper bound on $\|\tilde{S} - S\|$ over all perturbations satisfying (2). Now $(A+E)^{-1} = A^{-1} - \tilde{A}^{-1}EA^{-1}$. Hence on expanding the equation

$$\tilde{S} = C - (B + F)^{\mathrm{T}}(A + E)^{-1}(B + F),$$

we get

$$\tilde{S} = S + F^{\mathrm{T}}\tilde{A}^{-1}\tilde{B} - B^{\mathrm{T}}\tilde{A}^{-1}F + B^{\mathrm{T}}\tilde{A}^{-1}EA^{-1}B.$$

Taking norms and using the bounds (2), we get

$$\begin{aligned}\delta \leq \epsilon(&\|\tilde{A}^{-\frac{1}{2}}\tilde{B}\|\|\tilde{A}^{-\frac{1}{2}}\|\|B\| + \|\tilde{A}^{-\frac{1}{2}}B\|\|\tilde{A}^{-\frac{1}{2}}\|\|B\| \\ &+ \|A^{-\frac{1}{2}}B\|\|\tilde{A}^{-\frac{1}{2}}B\|\|A^{-\frac{1}{2}}\|\|\tilde{A}^{-\frac{1}{2}}\|\|A\|),\end{aligned}$$

It then follows from (5) that

$$\begin{aligned}\|\tilde{S} - S\| \le &\frac{\epsilon}{1-\kappa^{\frac{1}{2}}\epsilon}\bigl(\|\tilde{A}^{-\frac{1}{2}}\tilde{B}\|\|A^{-\frac{1}{2}}\|\|B\| + \|\tilde{A}^{-\frac{1}{2}}B\|\|A^{-\frac{1}{2}}\|\|B\| \\ &+ \kappa\|A^{-\frac{1}{2}}B\|\|\tilde{A}^{-\frac{1}{2}}B\|\bigr).\end{aligned} \tag{6}$$

We next show that

$$\|A^{-\frac{1}{2}}B\|^2, \|\tilde{A}^{-\frac{1}{2}}B\|^2, \|\tilde{A}^{-\frac{1}{2}}\tilde{B}\|^2 \le \|C\| + \delta. \tag{7}$$

To show the first inequality note that since $S \trianglerighteq 0$, we must have

$$\|C\| + \delta \ge \|C\| \ge \|B^{\mathrm{T}}A^{-1}B\| = \|A^{-\frac{1}{2}}B\|^2.$$

For the second inequality, take $F = 0$. Then $\|\tilde{S} - S\| \le \delta$. Since $S \trianglerighteq 0$, we have $\tilde{S} + \delta I \trianglerighteq 0$. Hence

$$\|C\| + \delta = \|C + \delta I\| \ge \|B\tilde{A}^{-1}B^{\mathrm{T}}\| = \|\tilde{A}^{-\frac{1}{2}}B^{\mathrm{T}}\|^2.$$

The third inequality follows similarly. Combining (6) and (7) we get

$$\delta \le \frac{\epsilon}{1-\kappa^{\frac{1}{2}}\epsilon}\bigl(2\|A^{-\frac{1}{2}}\|\|B\|(\|C\|+\delta)^{\frac{1}{2}} + \kappa(\|C\|+\delta)\bigr). \tag{8}$$

The next step is to bound $\|B\|$. Let x and y be vectors of norm one such that $\|B\| = y^{\mathrm{T}}Bx$. Then the matrix

$$\begin{pmatrix} x^{\mathrm{T}} & 0 \\ 0 & y^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} A & B^{\mathrm{T}} \\ B & C \end{pmatrix}\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} = \begin{pmatrix} x^{\mathrm{T}}Ax & x^{\mathrm{T}}B^{\mathrm{T}}y \\ y^{\mathrm{T}}Bx & y^{\mathrm{T}}Cy \end{pmatrix}$$

is positive semidefinite and has nonnegative determinant. It follows that

$$\|B\|^2 = (y^{\mathrm{T}}Bx)^2 < (x^{\mathrm{T}}Ax)(y^{\mathrm{T}}Cy) \le \|A\|\|C\|. \tag{9}$$

Combining (8) and (9), we get

$$\delta \le \epsilon(\|C\| + \delta)\beta.$$

If this inequality is solved for δ, the result is (4). ∎

The bound (4) is a normwise relative perturbation bound. However, the error is measured relative to $\|C\|$. If $\|S\|$ is comparable to $\|C\|$ (it cannot be bigger), then

$\kappa(A)$ controls the size of the relative perturbation in S. The bound deteriorates as $\|S\|$ approaches zero. This is to be expected, since if $\|S\|$ is very much smaller than $\|C\|$ the errors G in C can overwhelm S.

The structured bounds (2) allow for disparities in scaling: C, for example, may be much smaller than A [though by (9) $\|B\|$ must be bounded by the geometric mean of $\|A\|$ and $\|C\|$].

The factor $(1-\epsilon\beta)$ in (4) compensates for the fact that $C - B\tilde{A}^{-1}B^{\mathrm{T}}$ and $C - \tilde{B}\tilde{A}^{-1}B\tilde{B}$ need not be positive semidefinite. However if C is positive definite, then these two matrices are also positive definite for all sufficiently small ϵ, and in this case the factor $(1-\epsilon\beta)$ may be deleted. In any event, the asymptotic form of the bound is

$$\frac{\|\tilde{S}-S\|}{\|C\|} \lesssim \epsilon + (2\kappa^{\frac{1}{2}}(A) + \kappa(A))\epsilon \leq 3\kappa(A)\epsilon$$

References

[1] N.J. Higham. Analysis of the Cholesky decomposition of a semi-definite matrix. In M.G. Cox and S.J. Hammarling, editors, *Reliable Numerical Computation*, pages 161–185, Oxford, 1990. Clarenden Press.

[2] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, Boston, 1990.